

# Option Pricing in the Moderate Deviations Regime

Peter Friz  
TU and WIAS Berlin  
friz@math.tu-berlin.de

Stefan Gerhold  
TU Wien  
sgerhold@fam.tuwien.ac.at

Arpad Pinter\*  
TU Wien  
apinter@fam.tuwien.ac.at

April 6, 2016

## Abstract

We consider call option prices in diffusion models close to expiry, in an asymptotic regime (“moderately out of the money”) that interpolates between the well-studied cases of at-the-money options and out-of-the-money fixed-strike options. First and higher order small-time moderate deviation estimates of call prices and implied volatility are obtained. The expansions involve only simple expressions of the model parameters, and we show in detail how to calculate them for generic local and stochastic volatility models. Some numerical examples for the Heston model illustrate the accuracy of our results.

## 1 Introduction

Small-maturity approximations of option prices have been studied extensively in recent years. While out-of-the-money calls with fixed strike and at-the-money calls have both been thoroughly investigated, there is a significant asymptotic regime lying between the two. It has received little attention, and, to the best of our knowledge, none at all in the classical diffusion case. The aim of the present paper is to fill this gap. The “moderately out-of-the-money” regime reflects the reality of quoted option prices (strikes move closer to the money as expiry shrinks), while offering excellent analytic tractability.

To put our results into perspective, we recall some well-known facts on option price approximations close to expiry. We write  $C = C(K, t)$  for the call price, and  $c(k, t)$  if we wish to express it as a function of log-moneyness  $k$ :

$$C(S_0 e^k, t)/S_0 = c(k, t). \quad (1.1)$$

We start with the **at-the-money** (short: **ATM**) regime  $k = 0$ . In the Black-Scholes model, writing  $c(k, t) = c_{\text{BS}}(0, t; \sigma)$  with volatility parameter  $\sigma > 0$ , we have the following ATM call price behaviour,

$$c_{\text{BS}}(0, t; \sigma) \sim \frac{\sigma\sqrt{t}}{\sqrt{2\pi}}, \quad t \downarrow 0.$$

---

\*We gratefully acknowledge financial support from DFG grant FR2943/2 (P. Friz) resp. the Austrian Science Fund (FWF) under grant P 24880 (S. Gerhold, A. Pinter). This work was first presented by one of us (Friz) at Global Derivates 2015 in Amsterdam, participants are thanked for their feedback.

The same is actually true [31] in a generic semimartingale model with diffusive component (with spot volatility  $\sigma_0 = \sqrt{v_0} > 0$ ):

$$c(0, t) \sim \frac{\sigma_0 \sqrt{t}}{\sqrt{2\pi}}, \quad t \downarrow 0, \quad (1.2)$$

and this translates to the generic ATM implied volatility formula (even in presence of jumps, as long as  $v_0 > 0$ )

$$\sigma_{\text{imp}}^2(0, t) = v_0 + o(1), \quad t \downarrow 0.$$

Higher order terms in  $t$  will be model dependent. For instance, in the Heston case, with variance dynamics  $dV = -\kappa(V - \bar{v})dt + \eta\sqrt{V}dW$ , implied volatility has the following ATM expansion:

$$\begin{aligned} \sigma_{\text{imp}}^2(0, t) &= v_0 + a(0)t + o(t), \\ a(0) &= -\frac{\eta^2}{12} \left(1 - \frac{\rho^2}{4}\right) + \frac{v_0 \rho \eta}{4} + \frac{\kappa}{2} (\bar{v} - v_0). \end{aligned} \quad (1.3)$$

This is Corollary 4.4 in [16], and we note that  $a(0)$  has no easy interpretation in terms of the model parameters.

Relaxing  $k = 0$  to  $k = o(\sqrt{t})$  amounts to what we dub “**almost-ATM**” (short: **AATM**) regime.<sup>1</sup> (In particular,  $k \approx t^\beta$  is in the AATM regime if and only if  $\beta > 1/2$ .) Again for generic semimartingale models with diffusive component and spot volatility  $\sigma_0 > 0$ , it is easy to see [6, 31] that the ATM asymptotics (1.2) imply the following almost-ATM asymptotics:

$$c(k_t, t) \sim \frac{\sigma_0 \sqrt{t}}{\sqrt{2\pi}}, \quad k_t = o(\sqrt{t}), \quad t \downarrow 0.$$

This fails when  $k_t$  ceases to be  $o(\sqrt{t})$ . Indeed, for  $k_t = \theta\sqrt{t}$  with constant factor  $\theta > 0$ , we have [6, 31]:

$$c(k_t, t) \sim \mathbb{E}[N(-\theta, \sigma_0^2)^+] \sqrt{t}.$$

This, too, holds true in the stated semimartingale generality. In any case, the proof is based on the Lévy case with non-zero diffusivity  $v_0$ , and the result follows from comparison results which imply that the difference is negligible to first order. For a thorough discussion of the regime  $k = O(\sqrt{t})$  in the (local) diffusion case, see [34].

Beyond this regime, call price asymptotics change considerably. For instance, take an additional slowly diverging factor  $\log(1/t)$ ,

$$k_t = \theta \sqrt{t \log(1/t)}.$$

Even in the Black-Scholes model, we now loose the  $\sqrt{t}$ -behaviour of call prices seen above and in fact

$$c_{\text{BS}}(0, t; \sigma) = t^{\frac{1}{2} + \frac{\theta^2}{2\sigma^2}} \ell(t),$$

for some slowly varying function  $\ell(t)$ , see [30]. On the other hand, in a genuine **out-of-the-money** (short: **OTM**) situation, with  $k_t \equiv k > 0$  fixed, option values are exponentially small in diffusion models, and we are in the realm of *large deviation theory*. For instance,

$$c_{\text{BS}}(k, t; \sigma) \approx \exp\left(-\frac{\Lambda_{\text{BS}}(k)}{t}\right), \quad k > 0 \text{ fixed}, \quad t \downarrow 0,$$

Table 1: Asymptotic behavior of short-maturity call options,  $t \downarrow 0$ 

Process Type	ATM (at-the-money)  $K = S_0$	AATM (almost at-the-money)  $\log \frac{K}{S_0} \approx (\text{const})t^\beta$ $\beta > 1/2$	MOTM (moderately out-of-the-money)  $\log \frac{K}{S_0} \approx (\text{const})t^\beta$ $0 < \beta < 1/2$	OTM (out-of-the-money)  $\log(K/S_0) \equiv k > 0$
Black-Scholes	$O(\sqrt{t})$ , elementary	$O(\sqrt{t})$ , elementary	$\exp\left(-\frac{\text{const}}{t^{1-2\beta}}\right)$	$\exp\left(-\frac{\text{const}}{t}\right)$ , elementary
Stochastic volatility (diffusion model)	$O(\sqrt{t})$ , [7, 31]	$O(\sqrt{t})$ , [31]	$\exp\left(-\frac{\text{const}}{t^{1-2\beta}}\right)$	$\exp\left(-\frac{\text{const}}{t}\right)$ , [3]
Jump diffusion/ general semi- martingale with diff. component	$O(\sqrt{t})$ , [31]	$O(\sqrt{t})$ , [31]	$O(t)$ in Lévy models, [30]	$O(t)$ , [1, 7]

with  $\Lambda_{\text{BS}}(k) = \frac{1}{2}k^2/\sigma^2$  in the Black-Scholes model.<sup>2</sup> Similar results appear in the literature, with different level of mathematical rigor, for other and/or generic diffusion models [2, 7, 14, 35].

Throughout the paper, we reserve the term out-of-the-money (OTM) for *fixed* OTM log-strike  $k > 0$ , to distinguish this regime from the *moderately* out-of-the-money regime that we now introduce. Our basic observation is that for

$$k \approx (\text{const})t^\beta \quad (1.4)$$

the cases of  $\beta > \frac{1}{2}$ , resp.  $\beta = 0$ , are covered by the afore-discussed AATM, resp. OTM, results. This leaves open a significant gap, namely  $\beta \in (0, \frac{1}{2})$ , which we call **moderately out-of-the-money** (short: **MOTM**). We have a threefold interest in this MOTM regime,

$$k \approx (\text{const})t^\beta \quad \text{for } \beta \in (0, \tfrac{1}{2}). \quad (1.5)$$

(i) First, it is very much related to the *reality of quoted (short-dated) option prices*, where strikes of option price data with acceptable bid-ask spreads tend to accumulate “around the money”, as illustrated in Figure 1. It is then very natural to analyse strikes  $k = O(t^\beta)$  for some  $\beta > 0$ , and there is no reason why quoted strikes should always be almost-ATM, which effectively means an extreme concentration around the money thanks to  $\beta > 1/2$ .

(ii) The second reason is *mathematical convenience*. In contrast to the genuine OTM regime (large deviation regime) in which the rate function  $\Lambda(k)$  is notoriously difficult to analyse – often related to geodesic distance problems – MOTM naturally comes with a quadratic rate function and, most remarkably, higher order expansions are always explicitly computable in terms of the model parameters. The terminology *moderately*-OTM (MOTM) is in fact in reference to *moderate deviations theory*, which effectively interpolates between the central limit and large deviations regimes. (This also identifies the AATM regime as central limit regime, where asymptotics are precisely those of the Black-Scholes model, which in turn is the rescaled Gaussian (in log-coordinates) limit of a general semimartingale model with diffusive component.)

(iii) Our third point is that MOTM expansions naturally involve quantities very familiar to practitioners, notably spot (implied) volatility, implied volatility skew and so on.

<sup>1</sup>The term “almost-ATM” seems new, but this regime was considered by a number of authors including [6, 31].

<sup>2</sup>More precisely,  $t \log c_{\text{BS}}(k, t; \sigma) \sim -\Lambda_{\text{BS}}(k)$ ,  $k > 0$  fixed,  $t \downarrow 0$ .

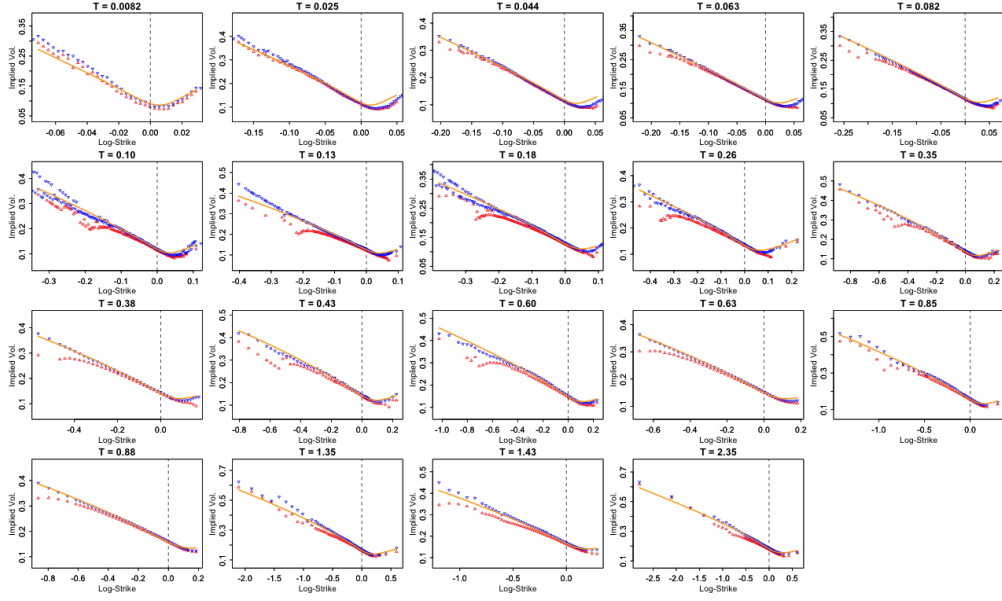


Figure 1: SPX volatility smiles as of 14-Aug-2013 (courtesy of J. Gatheral). Strikes of options with small remaining time to maturity ( $T = 0.0082$ ) are about  $e^{0.02} - 1 \approx 2\%$  around the money (spot), good data for a later time  $T = 0.26$  already has a range of  $\approx 30\%$ , the highest maturity  $T = 2.35$  has a range of  $\approx 65\%$  around the money.

In the Black-Scholes model, it is easy to check that we have the following MOTM asymptotics,

$$c_{BS}(k_t \equiv \theta t^\beta, t; \sigma) = \exp\left(-\frac{\theta^2}{2\sigma^2 t^{1-2\beta}}(1 + o(1))\right), \quad t \downarrow 0.$$

Loosely speaking, our main results assert that such relations (even of higher order) are true in great generality for diffusion models, all quantities are computable and then related to implied volatility expansions. We note in passing that, for Lévy models, the regime (1.5) has been studied in [30]; then, call prices decay algebraically rather than exponentially. For recent related results on *fractional* stochastic volatility models, see [17, 23]. Guillin’s work [24] on small-noise moderate deviations of diffusions should also be mentioned here; however, there the dynamics depend on a “fast” random environment (with motivation from physics, and no obvious financial interpretation), and also the non-degeneracy assumption (D) from [24] is not satisfied in our context.

The rest of the paper is organized as follows. Section 2 contains our main results, which translate asymptotics for the transition density of the underlying into MOTM call price asymptotics. The corresponding proofs are presented in Section 3. Section 4 gives the implied volatility expansion resulting from our call price approximations. Section 5 applies our main results to standard examples, namely generic local volatility models (Subsection 5.1), generic stochastic volatility models (Subsection 5.2), and the Heston model (Subsection 5.3). (As usual, the square-root degeneracy of the Heston model makes it difficult to apply results for general stochastic volatility models, so we verify the validity of our results – if formally applied to Heston – by a direct “affine” analysis.) Finally, in Section 6 we present a second approach at MOTM estimates, which employs the Gärtner-Ellis theorem from large deviation theory. Throughout we take zero rates,

natural in view of our short-time consideration. Also, w.l.o.g. we normalize spot to  $S_0 = 1$ .

## 2 MOTM option prices via density asymptotics

We consider a *general stochastic volatility* model, i.e. a positive martingale  $(S_t)_{t \geq 0}$  with dynamics

$$dS_t = S_t \sigma(t, \omega) dW_t$$

started (w.l.o.g.) at  $S_0 = 1$ . We assume that the stochastic volatility (process)  $\sigma$  itself is an Itô-diffusion, started at some deterministic value  $\sigma(0, \omega) \equiv \sigma_0$ , called *spot volatility*. Recall that in any such stochastic volatility model, the local (or effective) volatility is defined by

$$\sigma_{\text{loc}}^2(t, K) := \mathbb{E}[\sigma^2(t, \omega) | S_t = K].$$

As is well-known, the equivalent local volatility model

$$d\tilde{S}_t = \tilde{S}_t \sigma_{\text{loc}}^2(t, \tilde{S}_t) dW$$

has the property that  $\tilde{S}_t = S_t$  (in law) for all fixed times. See [5] for some precise technical conditions under which this holds true.<sup>3</sup> In particular then, European option prices  $C = C(K, t)$  match in both models. Recall also Dupire's formula in this context:

$$\sigma_{\text{loc}}^2(K, t) = \frac{\partial_t C(K, t)}{\frac{1}{2} K^2 \partial_{KK} C(K, t)}, \quad t > 0, K > 0. \quad (2.1)$$

We now state our two crucial conditions.

**Assumption 2.1** *For all  $t > 0$ ,  $S_t$  has a continuous pdf  $K \mapsto q(K, t)$ , which behaves asymptotically as follows for small time:*

$$q(K, t) \sim e^{-\frac{\Lambda(k)}{t}} t^{-1/2} \gamma(k), \quad t \downarrow 0, \quad (2.2)$$

*uniformly for  $K = e^k$  in some neighbourhood of  $S_0 = 1$ . The energy function  $\Lambda$  is smooth in some neighbourhood of zero, with  $\Lambda(0) = \Lambda'(0) = 0$ . Moreover,  $\lim_{k \rightarrow 0} \gamma(k) = \gamma(0) > 0$ .*

**Assumption 2.2** *For  $t \downarrow 0$  and  $K \rightarrow S_0 = 1$ , the local volatility function of  $(S_t)_{t \geq 0}$  converges to spot volatility:*

$$\lim_{\substack{K \rightarrow S_0 \\ t \downarrow 0}} \sigma_{\text{loc}}(K, t) = \sigma_0.$$

The latter assumption is (in diffusion models) fairly harmless. The first assumption is potentially (very) difficult to check, but fortunately we can rely on substantial recent progress in this direction [10, 11, 33]. We shall see in Section 5.2 in detail that both assumptions indeed hold in generic stochastic volatility models. Let us also note the fundamental relation between spot-volatility  $\sigma_0$  (actually equal to implied spot vol  $\sigma_{\text{imp}}(0, 0)$  here) and the Hessian of the energy function  $\Lambda = \Lambda(k)$ ,

$$\sigma_0 = \Lambda''(0)^{-1/2}.$$

(This is well-known (e.g. [12]) and also follows from Proposition 2.4 below.) Now we state our main result. We slightly generalize the log-strikes considered in (1.5), replacing the constant by an arbitrary slowly varying function  $\ell$ .

---

<sup>3</sup>The situation is very different with jumps, see [19].

**Theorem 2.3** *Under Assumptions 2.1 and 2.2, consider a moderately out-of-the-money call, in the sense that log-strike is*

$$k_t = t^\beta \ell(t), \quad t > 0, \quad (2.3)$$

where  $\ell > 0$  varies slowly at zero and  $\beta \in (0, \frac{1}{2})$ .

(i) *The call price satisfies the moderate deviation estimate*

$$\begin{aligned} c(k_t, t) &= \exp \left( -\frac{\Lambda''(0)}{2} \frac{k_t^2}{t} (1 + o(1)) \right) \\ &= \exp \left( -\frac{1}{2\sigma_0^2} \frac{k_t^2}{t} (1 + o(1)) \right), \quad t \downarrow 0. \end{aligned} \quad (2.4)$$

(ii) *If we restrict  $\beta$  to  $(0, \frac{1}{3})$ , then the following moderate second order expansion holds true:*

$$\begin{aligned} c(k_t, t) &= \exp \left( -\frac{1}{2} \Lambda''(0) \frac{k_t^2}{t} - \frac{1}{6} \Lambda'''(0) \frac{k_t^3}{t} (1 + o(1)) \right) \\ &= \exp \left( -\frac{1}{2\sigma_0^2} \frac{k_t^2}{t} \left( 1 - \frac{\mathcal{S}}{\sigma_0^2} k_t (1 + o(1)) \right) \right), \quad t \downarrow 0, \end{aligned} \quad (2.5)$$

with spot-variance  $\sigma_0^2$ , equal to  $\sigma_{\text{imp}}^2(0, 0)$ , and implied variance skew  $\mathcal{S} = \frac{\partial}{\partial k} \big|_{k=0} \sigma_{\text{imp}}^2(k, 0)$ .

In particular, for  $\ell \equiv \theta > 0$ , we have the (first order) expansion

$$t^{1-2\beta} \log c(\theta t^\beta, t) \sim -\frac{\theta^2}{2\sigma_0^2}, \quad t \downarrow 0,$$

exhibiting a quadratic rate function  $\theta \mapsto \theta^2/2\sigma_0^2$ , typical of *moderate deviation problems*.<sup>4</sup>

The quantities  $\Lambda''(0), \Lambda'''(0), \dots$  appearing above are *always computable* from the initial values and the diffusion coefficients of the stochastic volatility model. This is in stark contrast to the OTM regime, where one needs  $\Lambda(k)$ , which is in general not available in closed form (with some famous exceptions, like the SABR model). We quote the following result on  $N$ -factor models from Osajima [33], and refer to Section 5.2 for detailed calculations in a two-factor stochastic volatility model.

**Proposition 2.4** *Assume that  $(\log S, \sigma^1, \dots, \sigma^{N-1})$  is Markov, started at  $(0, \bar{\sigma}_0)$  with  $\bar{\sigma}_0 \in \mathbb{R}^{N-1}$  and  $\bar{\sigma}_0^1 > 0$ , with stochastic volatility  $\sigma \equiv \sigma^1$ , where the generator has (non-degenerate) principal part  $\sum a^{ij} \partial_{ij}$  in the sense that  $a^{-1}$  defines a Riemannian metric. Then*

$$\Lambda(k) = \frac{1}{2b_1} k^2 - \frac{b_2}{3b_1^3} k^3 + \left( -\frac{b_3}{4b_1^4} + \frac{b_2^2}{2b_1^5} \right) k^4 + O(k^5), \quad k \rightarrow 0,$$

where the coefficients are given by

$$\begin{aligned} b_1 &= \int_0^1 a^{11}(t, \bar{\sigma}_0) dt \\ b_2 &= \frac{3}{2} \int_0^1 (V a^{11})(t, \bar{\sigma}_0) dt \\ b_3 &= 2 \int_0^1 (V^2 a^{11})(t, \bar{\sigma}_0) dt + \frac{1}{2} \int_0^1 \Gamma(a^{11}, a^{11})(t, \bar{\sigma}_0) dt, \end{aligned}$$

---

<sup>4</sup>Recall that the MD rate function for a centered i.i.d. sequence  $(X_n)_{n \geq 1}$  is given by  $\theta \mapsto \theta^2/(2\text{Var}(X_1))$ . This is the “moderate” version of Cramér’s theorem, see Theorem 3.7.1 in [9].

using the functions

$$(Vf)(t, x) = \sum_{i=1}^N a^{1i}(t, x) \int_t^1 \frac{\partial f}{\partial x_i}(s, x) ds,$$

$$\Gamma(f, g)(t, x) = \sum_{i,j=1}^N a^{ij}(t, x) \left( \int_t^1 \frac{\partial f}{\partial x_i}(s, x) ds \right) \left( \int_t^1 \frac{\partial g}{\partial x_j}(s, x) ds \right).$$

**Proof.** See Osajima [33], Theorem 1(1), with  $T = 1$ . ■

The following result presents a higher-order expansion in the MOTM regime. It yields an asymptotically equivalent expression for call prices (and not just logarithmic asymptotics).

**Theorem 2.5** *Under the assumptions of Theorem 2.3, we set  $k_t = t^\beta \ell(t)$  with  $\beta \in (0, \frac{1}{2})$  and  $\ell > 0$  slowly varying at zero, as above. Then the logarithm of the call price has the following refined MOTM expansion:*

$$\log c(k_t, t) = - \sum_{m=2}^{\lfloor 1/\beta \rfloor} \frac{\Lambda^{(m)}(0)}{m!} \frac{k_t^m}{t} + \left(2\beta - \frac{3}{2}\right) \log \frac{1}{t} - 2 \log \ell(t) + \log(\gamma(0)v_0^2) + o(1), \quad t \downarrow 0. \quad (2.6)$$

Equivalently,

$$c(k_t, t) \sim \gamma(0)v_0^2 \frac{t^{3/2-2\beta}}{\ell(t)^2} \exp \left( - \sum_{m=2}^{\lfloor 1/\beta \rfloor} \frac{\Lambda^{(m)}(0)}{m!} \frac{k_t^m}{t} \right), \quad t \downarrow 0.$$

If  $1/\beta$  is not an integer, then  $k_t^m/t$  tends to infinity for  $m = \lfloor 1/\beta \rfloor$ , of order  $t^{\beta \lfloor 1/\beta \rfloor - 1}$  (up to a slowly varying factor). If  $1/\beta$  is an integer, on the other hand, then the last summand of the sum  $\sum_{m=2}^{\lfloor 1/\beta \rfloor}$  in (2.6) is of order  $\ell(t)$ , which means that the following term  $\log(1/t)$  may be asymptotically larger. The upper summation limit  $\lfloor 1/\beta \rfloor$  thus ensures that no irrelevant (i.e.,  $o(1)$ ) terms are contained in the sum. Note that we have  $\lfloor 1/\beta \rfloor = 2$  for  $\beta \in (\frac{1}{3}, \frac{1}{2})$ , and  $\lfloor 1/\beta \rfloor \geq 3$  for  $\beta \in (0, \frac{1}{3})$ , and so (2.6) is consistent with (2.4) resp. (2.5).

The passage from the (derivatives of the) energy to ATM derivatives of the implied volatility in the short time limit is best conducted via the Berestycki-Busca-Florent (short: BBF) formula that was proved in [2]. (That said, these relations are also a direct consequence of our expansions, as is pointed out in Section 4.) In this regard, we have

**Theorem 2.6** *Suppose that  $\Lambda$  is a function with the properties required in Assumption 2.1, with  $\Lambda''(0) = \sigma_0^{-2} = v_0^{-1}$ , and that the Berestycki-Busca-Florent formula  $\sigma_{\text{imp}}^2(0, k) = k^2/2\Lambda(k)$  holds. Then the small-time ATM implied variance skew and curvature, respectively, relate to  $\Lambda$  via*

$$\mathcal{S} := \frac{\partial}{\partial k} \Big|_{k=0} \sigma_{\text{imp}}^2(k, 0) = -\frac{1}{3} \frac{\Lambda'''(0)}{\Lambda''(0)^2} \quad (2.7)$$

and

$$\mathcal{C} := \frac{\partial^2}{\partial k^2} \Big|_{k=0} \sigma_{\text{imp}}^2(k, 0) = \frac{\frac{2}{3}\Lambda'''(0)^2 - \frac{1}{2}\Lambda''''(0)\Lambda''(0)}{3\Lambda''(0)^3}. \quad (2.8)$$

**Proof.** By the BBF formula and our smoothness assumptions on  $\Lambda$ ,

$$\begin{aligned}\sigma_{\text{imp}}^2(k, 0) &= \frac{k^2}{2\Lambda(k)} = k^2 \left( \Lambda''(0)k^2 + \frac{1}{3}\Lambda'''(0)k^3 + \frac{1}{12}\Lambda''''(0)k^4 + O(k^5) \right)^{-1} \\ &= \frac{1}{\Lambda''(0)} - \frac{1}{3} \frac{\Lambda'''(0)}{\Lambda''(0)^2} k + \left( \frac{\frac{1}{9}\Lambda'''(0)^2 - \frac{1}{12}\Lambda''''(0)\Lambda''(0)}{\Lambda''(0)^3} \right) k^2, \quad k \rightarrow 0.\end{aligned}$$

This implies (2.7) and (2.8). ■

**Remark 2.7** Proposition 2.4 combined with Theorem 2.6 allows to compute skew and curvature (and higher derivatives of the implied volatility smile, if desired) directly from the coefficients of a general stochastic volatility model. Related formulae for “general” (even non-Markovian) models also appear in the work of Durrleman (Theorem 3.1.1. in [12]; see also [13]). While not written in the setting of general Markovian diffusion models, and hence not in terms of the energy function  $\Lambda$ , they inevitably give the same results if applied to given parametric stochastic volatility models (see Section 3.1 in [12]). However, this work comes with some (seemingly) uncheckable assumptions, the drawback of which is discussed in Section 2.6 of [12].

### 3 Proofs of the main results

**Proof of Theorem 2.3.** Since the density of  $S_t$  satisfies  $q = \partial_{KK}C$ , we have, by Dupire’s formula (2.1),

$$C(K, t) = \int_0^t \partial_s C(K, s) ds = \int_0^t \frac{1}{2} q(K, s) K^2 \sigma_{\text{loc}}^2(K, s) ds.$$

Then, for  $K_t = e^{k_t}$  with  $k_t \downarrow 0$  as stated, we apply Assumption 2.2 as follows:

$$\begin{aligned}C(K_t, t) &= \int_0^t \frac{1}{2} q(K_t, s) K_t^2 \sigma_{\text{loc}}^2(K_t, s) ds \\ &\sim \frac{\sigma_0^2}{2} \int_0^t q(K_t, s) ds, \quad t \downarrow 0.\end{aligned}$$

And then, using local uniformity of our density expansion (2.2),

$$\begin{aligned}C(K_t, t) &\sim \frac{\sigma_0^2 \gamma(0)}{2} \int_0^t e^{-\frac{\Lambda(k_t)}{s}} s^{-1/2} dt \\ &= \frac{\sigma_0^2 \gamma(0)}{2} t \int_0^1 e^{-\frac{\Lambda(k_t)}{xt}} (xt)^{-1/2} dx \\ &= \frac{\sigma_0^2 \gamma(0)}{2} t^{1/2} \int_0^1 e^{-\frac{\Lambda(k_t)}{xt}} x^{-1/2} dx.\end{aligned}\tag{3.1}$$

Since  $\Lambda$  is smooth at zero, and using the fact that  $\Lambda(0) = \Lambda'(0) = 0$ , we have

$$\frac{\Lambda(k_t)}{t} \sim \frac{1}{2} \Lambda''(0) \frac{k_t^2}{t} \rightarrow \infty \quad \text{as } t \downarrow 0.$$

For small  $t$ , the integrand in (3.1) is thus concentrated near  $x = 1$ , and by the Laplace method (Theorem 7.1 in [32])

$$\int_0^1 e^{-\frac{\Lambda(k_t)}{xt}} x^{-1/2} dx \sim \frac{t}{\Lambda(k_t)} \exp\left(-\frac{\Lambda(k_t)}{t}\right).$$



Therefore,

$$\begin{aligned} C(K_t, t) &\sim \frac{\sigma_0^2 \gamma(0)}{2} \frac{t^{3/2}}{\Lambda(k_t)} \exp\left(-\frac{\Lambda(k_t)}{t}\right) \\ &\sim v_0^2 \gamma(0) \frac{t^{3/2}}{k_t^2} \exp\left(-\frac{\Lambda(k_t)}{t}\right), \quad t \downarrow 0, \end{aligned} \quad (3.2)$$

which implies (recall the notation  $c$  resp.  $C$  from (1.1))

$$\begin{aligned} -\log c(k_t, t) &= \frac{\Lambda(k_t)}{t} - \log \frac{t^{3/2}}{k_t^2} + O(1) \\ &= \frac{1}{t} \left( \frac{1}{2} \Lambda''(0) k_t^2 + \frac{1}{6} \Lambda'''(0) k_t^3 + O(k_t^4) \right) + O\left(\log \frac{k_t^2}{t^{3/2}}\right). \end{aligned} \quad (3.3)$$

to prove (i) and (ii), we thus need to argue that  $k_t^2/t$  dominates  $\log(k_t^2 t^{-3/2})$  if  $\beta \in (0, \frac{1}{2})$ , and that  $k_t^3/t$  dominates  $\log(k_t^2 t^{-3/2})$  if  $\beta \in (0, \frac{1}{3})$ . For  $m \in \{2, 3\}$ , we calculate

$$\begin{aligned} \frac{k_t^m/t}{|\log(k_t^2 t^{-3/2})|} &= \frac{t^{m\beta-1} \ell(t)^m}{|\log(t^{2\beta-3/2} \ell(t)^2)|} \\ &= \frac{t^{m\beta-1} \ell(t)^m}{|(2\beta - \frac{3}{2}) \log t + 2 \log \ell(t)|}. \end{aligned}$$

From Proposition 1.3.6 (i) in [4] we know that  $\log \ell(t) = o(\log t)$ , and so

$$\frac{k_t^m/t}{|\log(k_t^2 t^{-3/2})|} \sim \frac{t^{m\beta-1} \ell(t)^m}{|(2\beta - \frac{3}{2}) \log t|}, \quad t \downarrow 0.$$

This tends to infinity for  $m = 2$  and  $\beta \in (0, \frac{1}{2})$ , and for  $m = 3$  and  $\beta \in (0, \frac{1}{3})$ , as desired. ■

Inspecting the preceding proof, it is easy to see that we can expand  $\log c(k_t, t)$  further:

**Proof of Theorem 2.5.** Taking logs in (3.2) yields

$$\log c(k_t, t) = -\frac{\Lambda(k_t)}{t} + \log \frac{t^{3/2}}{k_t^2} + \log(\gamma(0) v_0^2) + o(1).$$

Then (2.6) follows by Taylor expanding  $\Lambda$ . Note that  $k_t^m/t = o(1)$  for  $m \geq \lfloor 1/\beta \rfloor + 1$ . ■

## 4 Implied volatility

We now investigate some relations of MOTM call price expansions with implied volatility.

**Corollary 4.1** *Under the assumptions of Theorem 2.3, put  $k_t = t^\beta \ell(t)$  with  $\beta \in (0, \frac{1}{3})$  and  $\ell > 0$  slowly varying. Then the implied volatility has the following MOTM expansion:*

$$\sigma_{\text{imp}}(k_t, t) = \sigma_0 - \frac{1}{6} \sigma_0^3 \Lambda'''(0) k_t (1 + o(1)) \quad (4.1)$$

**Proof.** We use our main result (Theorem 2.3) in conjunction with a transfer result of Gao and Lee [20]. As the call price tends to zero, we are in case “(–)” of [20] (defined on p. 354 of

that paper). The notation  $L, V$  of [20] means  $L = -\log c(k_t, t)$  resp.  $V = t^{1/2}\sigma_{\text{imp}}(k_t, t)$ , the dimensionless implied volatility. Then Corollary 7.2 of [20] implies that

$$V = \frac{k_t}{\sqrt{2L}}(1 + O(t^{1-2\beta-\varepsilon})) + O(t^{5/2-4\beta-\varepsilon}), \quad t \downarrow 0. \quad (4.2)$$

Here,  $\varepsilon > 0$  denotes an arbitrarily small constant that serves to eat up slowly varying functions in  $O$ -estimates. By part (ii) of Theorem 2.3, we have

$$2L = \frac{1}{\sigma_0^2} \frac{k_t^2}{t} + \frac{\Lambda'''(0)}{3} \frac{k_t^3}{t} (1 + o(1)).$$

Inserting this into (4.2) gives

$$\sigma_{\text{imp}}(k_t, t) = t^{-1/2} k_t \left( \sigma_0 \frac{t^{1/2}}{k_t} - \frac{\sigma_0^3 \Lambda'''(0)}{6} t^{1/2} (1 + o(1)) \right) + O(t^{2-4\beta-\varepsilon}),$$

which yields (4.1). ■

The above corollary has some interesting consequences. Under the sheer assumption that implied volatility has a first order Taylor expansion for small maturity and small log-strike of the form

$$\sigma_{\text{imp}}(k, t) = \sigma_0 + \partial_k \sigma_{\text{imp}}(0, 0) k + o(k) + O(t), \quad t \downarrow 0, \quad k = o(1), \quad (4.3)$$

then of course in the MOTM regime, we have  $t \ll k_t$ , and so the  $k$ -term dominates the  $O(t)$ -term, which in turn identifies the implied variance skew as

$$\mathcal{S} = \lim_{t \downarrow 0} \frac{2\sigma_0}{k_t} (\sigma_{\text{imp}}(k_t, t) - \sigma_0). \quad (4.4)$$

On the other hand, Corollary 4.1 now implies that the right-hand side of (4.4) equals  $-\frac{1}{3}\sigma_0^4 \Lambda'''(0)$ . We have thus arrived at an alternative proof of the skew representation (2.7) in terms of the energy function, without using the BBF formula. (The curvature and higher order derivatives of the ATM smile can be dealt with similarly, if desired.)

## 5 Examples

### 5.1 Generic local volatility models

Clearly, Assumption 2.2 is satisfied for any local volatility model, assuming continuity of the local volatility function. We now discuss Assumption 2.1, and show how to compute our MOTM expansions. First consider the time-homogeneous local volatility model

$$dS_t = \sigma(S_t) S_t dW_t, \quad S_0 = 1, \quad (5.1)$$

where  $\sigma$  is  $C^2$  on  $(0, \infty)$ . An expansion of the pdf  $q(\cdot, t)$  of  $S_t$  has been worked out in Gatheral et al. [22]. They assume growth conditions on  $\sigma$  and its derivatives, which can be alleviated by the principle of not feeling the boundary (Appendix A of [22]). Proposition 2.1 of [22] says that

$$q(e^k, t) \sim \frac{e^{-k} u_0(0, k)}{\sqrt{2\pi t}} \exp\left(-\frac{\Lambda(k)}{t}\right), \quad t \downarrow 0,$$

uniformly in  $k$ , where the energy function is given by (cf. Varadhan [37])

$$\Lambda(k) = \frac{1}{2} \left( \int_0^k \frac{dx}{\sigma(e^x)} \right)^2,$$

and

$$u_0(0, k) = \sigma(1)^{1/2} \sigma(e^k)^{-3/2} e^{-k/2}.$$

(Recall that we normalize spot to  $S_0 = 1$  throughout.) This shows that Assumption 2.1 is satisfied, with  $\gamma(0) = 1/(\sqrt{2\pi}\sigma(1))$ . To evaluate the expansions from Theorem 2.3, we compute the derivatives of  $\Lambda$ :

$$\Lambda'(k) = \frac{1}{\sigma(e^k)} \int_0^k \frac{dx}{\sigma(e^x)}, \quad \Lambda''(k) = \frac{1}{\sigma(e^k)^2} - \frac{e^k \sigma'(e^k)}{\sigma(e^k)^2} \int_0^k \frac{dx}{\sigma(e^x)},$$

and

$$\Lambda'''(k) = -\frac{3e^k \sigma(e^k)}{\sigma(e^k)^3} + \left( \frac{2e^{2k} \sigma'(e^k)^2}{\sigma(e^k)^3} - \frac{\sigma(e^k)''}{\sigma(e^k)^2} \right) \int_0^k \frac{dx}{\sigma(e^x)},$$

which yield

$$\Lambda''(0) = \frac{1}{\sigma(1)^2} = \frac{1}{\sigma(S_0)^2}$$

and

$$\Lambda'''(0) = -\frac{3\sigma'(1)}{\sigma(1)^3} = -\frac{3\sigma'(S_0)}{\sigma(S_0)^3}. \quad (5.2)$$

Alternatively, these expressions can be obtained from Proposition 2.4. Since the assumptions of Theorem 2.3 are satisfied, we obtain the following MOTM call price estimates, where  $k_t = \theta t^\beta$  and  $\theta > 0$ :

$$\begin{aligned} c(k_t, t) &= \exp \left( -\frac{\theta^2}{2\sigma(1)^2 t^{1-2\beta}} (1 + o(1)) \right), \quad \beta \in (0, \tfrac{1}{2}), \quad t \downarrow 0, \\ c(k_t, t) &= \exp \left( -\frac{\theta^2}{2\sigma(1)^2 t^{1-2\beta}} - \frac{\sigma'(1)}{2\sigma(1)^3} \frac{\theta^3}{t^{1-3\beta}} (1 + o(1)) \right), \quad \beta \in (0, \tfrac{1}{3}), \quad t \downarrow 0. \end{aligned}$$

Recall that we denote by  $\mathcal{S}$  the (limiting small-time ATM) implied *variance* skew, and so the implied *volatility* skew is given by  $\mathcal{S}/2\sigma_0$ , which equals  $\mathcal{S}/2\sigma(1) = \mathcal{S}/2\sigma(S_0)$  in the model (5.1). From (2.7) and (5.2) we recover that the *local* skew  $\sigma'(1) = \sigma'(S_0)$  equals twice the implied volatility skew,

$$\sigma'(S_0) = 2 \times \frac{\mathcal{S}}{2\sigma(S_0)},$$

as observed by Henry-Labordère (Remark 5.2 in [26]). Generic time-inhomogeneous local volatility models

$$dS_t = \sigma(S_t, t) S_t dW_t$$

could be treated very similarly, using the heat kernel expansion in Section 3 of [22], itself taken from Yosida [38].

## 5.2 Generic stochastic volatility models

We now discuss the results of Section 2 in generic stochastic volatility models. Rigorous conditions under which stochastic volatility models satisfy Assumption 2.1 can be found in [10, 33]. The function  $\Lambda$  is given by the Riemannian metric associated to the model:  $2\Lambda(k)$  is the squared geodesic distance from  $(S_0 = 1, \sigma_0)$  to  $\{(K, \sigma) : \sigma > 0\}$  with  $K = e^k$ . Theorem 2.2 in Berestycki, Busca, and Florent [3] gives conditions under which Assumption 2.2, concerning convergence of local volatility, is true.

Now we describe how the expressions appearing in the expansions from Theorem 2.3 can be computed explicitly in a generic two-factor stochastic volatility model

$$\begin{aligned} dS_t &= S_t \sqrt{V_t} dW_t, \quad S_0 = 1, \\ dV_t &= (\dots) dt + \eta \sqrt{V_t} \nu(V_t) dZ_t, \quad V_0 = v_0 > 0, \end{aligned} \quad (5.3)$$

where  $\nu: \mathbb{R} \rightarrow \mathbb{R}$  and  $d\langle W, Z \rangle_t = \rho dt$ . The Heston model ( $\nu(v) \equiv 1$ ) and the 3/2-model ( $\nu(v) = v$ ; see [29]) are special cases. The infinitesimal generator  $L$  of the stochastic process  $(S, V)$ , neglecting first order terms, can be written as

$$Lf \approx \frac{1}{2} \text{Tr} \left( \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} D^2 f \right), \quad f \in C^2(\mathbb{R}^2),$$

where  $D^2 f$  denotes the Hessian matrix of  $f$ , and the coefficient matrix  $g = (g_{ij})$  in this model is given by

$$g = \begin{pmatrix} v & \rho \eta v \nu(v) \\ \rho \eta v \nu(v) & \eta^2 v \nu(v)^2 \end{pmatrix}.$$

We define the constants  $b_1 = g_{11}|_{v=v_0} = v_0$  and  $b_2 = \frac{3}{4} \sum_{i=1}^2 g_{1i} \partial_i g_{11}|_{v=v_0} = \frac{3}{4} \rho \eta v_0 \nu(v_0)$ . If we assume that the coefficients in (5.3) are nice enough to justify application of the (marginal) density expansion obtained in [10] or part (2) of Theorem 1 in [33], we get the desired small-time density expansion (2.2). Moreover, thanks to Proposition 2.4,

$$\Lambda(k) = \frac{1}{2b_1} k^2 - \frac{b_2}{3b_1^3} k^3 + O(k^4)$$

in a neighborhood of 0. Therefore the quantities  $\Lambda''(0) = v_0^{-1} = \sigma_0^{-2}$  and  $\Lambda'''(0) = -\frac{3}{2} \rho \eta \nu(v_0)/v_0^2$  can easily be computed, as well as the small-time ATM implied variance skew

$$\mathcal{S} = -\frac{v_0^2}{3} \Lambda'''(0) = \frac{\rho \eta}{2} \nu(v_0).$$

Thus, all quantities appearing in our expansions (Theorem 2.3, Corollary 4.1) have very simple expressions in terms of the model parameters.

## 5.3 The Heston model

This section contains an application of the results of Sections 2 and 4 to the familiar case of the Heston model, where many explicit “affine” computations are possible. The Heston model is not in the scope of the general results implying Assumptions 2.1 and 2.2, which we recalled at the beginning of Section 5.2. We will explain how both assumptions can be verified rigorously by a dedicated analysis; full details would involve rather dull repetition of arguments that are found in the literature in a very similar form, and are therefore omitted. The model dynamics are

$$\begin{aligned} dS_t &= S_t \sqrt{V_t} dW_t, \quad S_0 = 1, \\ dV_t &= -\kappa (V - \bar{v}) dt + \eta \sqrt{V_t} dZ_t, \quad V_0 = v_0 > 0, \end{aligned}$$

where  $\bar{v}, \kappa, \eta > 0$ , and  $d\langle W, Z \rangle_t = \rho dt$  with  $\rho \in (-1, 1)$ . According to Forde and Jacquier [14], the first order OTM (large deviations) behavior of the call prices is

$$t \log c_{\text{He}}(k, t) \sim -\Lambda_{\text{He}}(k), \quad k > 0 \text{ fixed, } t \downarrow 0, \quad (5.4)$$

where  $\Lambda_{\text{He}}$  is the (not explicitly available) Legendre transform of

$$\begin{aligned} \Gamma(p) &= \frac{v_0 p}{\eta(\bar{\rho} \cot(\eta \bar{\rho} p/2) - \rho)} \\ &= \frac{v_0 p}{\eta(\bar{\rho}(\frac{1}{\eta \bar{\rho} p/2} + O(p)) - \rho)} \\ &= \frac{v_0 p}{\frac{1}{p/2} - \eta \rho + O(p)} \\ &= \frac{v_0 p^2/2}{1 - p\eta\rho/2 + O(p^2)} \\ &= \frac{v_0 p^2}{2} (1 + p\eta\rho/2 + O(p^2)). \end{aligned} \quad (5.5)$$

(We use the standard notation  $\bar{\rho}^2 = 1 - \rho^2$ .) This expansion implies

$$\Gamma''(0) = v_0 = \sigma_0^2. \quad (5.6)$$

The locally uniform density asymptotics (2.2) hold, as seen from an easy modification of the arguments in Forde, Jacquier, and Lee [16]. There, the Fourier representation of the call price was analysed by the saddle point method to obtain a refinement of (5.4). Proceeding completely analogously for the Fourier representation of the pdf of  $S_t$ , we get the density approximation

$$\begin{aligned} q_{\text{He}}(e^k, t) &= \frac{e^{-k}}{2\pi t} \int_{-\infty - ip^*(k)}^{\infty - ip^*(k)} \text{Re}(e^{iku/t} \phi_t(-u/t)) du \\ &= \exp\left(-\frac{\Lambda_{\text{He}}(k)}{t}\right) \frac{U(p^*(k))}{\sqrt{2\pi\Gamma''(k)}} t^{-1/2} (1 + o(1)), \quad t \downarrow 0, \end{aligned}$$

locally uniformly in  $k$ , where  $\phi_t$  is the characteristic function of  $X_t = \log S_t$ , and  $p^*$  and  $U$  are defined on p. 693 of [16]. (Note that [16] uses the notation  $\Lambda, \Lambda^*$  instead of our  $\Gamma, \Lambda_{\text{He}}$ .) From (5.6) and the fact that  $U(p^*(0)) = U(0) = 1$ , we see that the factor  $\gamma(k)$  from (2.2) converges to

$$\gamma(0) = \frac{1}{\sqrt{2\pi}\sigma_0} \quad (5.7)$$

as  $k \rightarrow 0$ .

To verify Assumption 2.2 (convergence of local volatility), the Dupire formula (2.1) can be subjected to an analysis similar to [8, 18]. More precisely,  $\partial_{KK} C(K, t)$  in the numerator of (2.1) is the pdf of  $S_t$ , the analysis of which we have just described. Virtually the same saddle point approach can be applied to the numerator  $\partial_t C(K, t)$ , yielding convergence of the quotient to  $\sigma_0^2$ .

We now calculate our MOTM asymptotic expansions for the Heston model. The Legendre transform  $\Lambda_{\text{He}}$  is given by  $\Lambda_{\text{He}}(k) = \sup_x \{kx - \Gamma(x)\}$  with maximizer  $x^* = x^*(k)$ . From general facts on Legendre transforms,

$$\Lambda''_{\text{He}}(k) = \frac{1}{\Gamma''(x^*(k))}.$$

Since  $x^*(0) = 0$ , we have

$$\Lambda''_{\text{He}}(0) = \frac{1}{\Gamma''(0)} = \frac{1}{v_0}.$$

From Theorem 2.3, with  $k_t = \theta t^\beta$  and  $\theta > 0$ , we then obtain the MOTM call price estimate

$$c_{\text{He}}(k_t, t) = \exp \left( -\frac{\theta^2}{2v_0 t^{1-2\beta}} (1 + o(1)) \right), \quad t \downarrow 0. \quad (5.8)$$

As for the second order expansion, from the expansion (5.5) of  $\Gamma$  we clearly see that

$$\Gamma'''(0) = \frac{3}{2} v_0 \eta \rho.$$

On the other hand, a general Legendre computation gives

$$\Lambda'''_{\text{He}}(k) = - \left( \frac{1}{\Gamma''(x^*(k))} \right)^2 \Gamma'''(x^*(k)) (x^*)'(k) = -(\Lambda''_{\text{He}}(k))^3 \Gamma'''(x^*(k)).$$

Therefore,

$$\Lambda'''_{\text{He}}(0) = -\frac{3}{2} \frac{\eta \rho}{v_0^2},$$

in accordance with the expression for generic two-factor models, found in Section 5.2. For  $\beta \in (0, \frac{1}{3})$ , Theorem 2.3 (ii) thus implies the second order expansion

$$c_{\text{He}}(k_t, t) = \exp \left( -\frac{\theta^2}{2v_0 t^{1-2\beta}} + \frac{\eta \rho}{4v_0^2} \frac{\theta^3}{t^{1-3\beta}} (1 + o(1)) \right), \quad t \downarrow 0. \quad (5.9)$$

By Theorem 2.5 and (5.7), we obtain the following refined call price expansions:

$$\log c_{\text{He}}(k_t, t) = -\frac{1}{2\sigma_0^2} \frac{k_t^2}{t} + \left( \frac{3}{2} - 2\beta \right) \log t + \log \frac{\sigma_0^3}{\sqrt{2\pi}} + o(1), \quad \beta \in \left( \frac{1}{3}, \frac{1}{2} \right), \quad (5.10)$$

$$\log c_{\text{He}}(k_t, t) = -\frac{1}{2\sigma_0^2} \frac{k_t^2}{t} + \frac{\eta \rho}{4v_0^2} \frac{k_t^3}{t} + \left( \frac{3}{2} - 2\beta \right) \log t + \log \frac{\sigma_0^3}{\sqrt{2\pi}} + o(1), \quad \beta \in \left( \frac{1}{4}, \frac{1}{3} \right). \quad (5.11)$$

From the relation (2.7) between implied variance skew and  $\Lambda'''(0)$ , we obtain the explicit expression  $\mathcal{S}_{\text{He}} = \eta \rho / 2$  for the skew. This agrees with Gatheral [21], p. 35. The implied volatility expansion (4.4) becomes

$$\sigma_{\text{imp}}(k_t, t) = \sigma_0 + \frac{\eta \rho}{4\sigma_0} k_t (1 + o(1)), \quad t \downarrow 0. \quad (5.12)$$

Figure 2 shows a good fit of this approximation, even for maturities that are not very small.

## 6 MOTM option prices via the Gärtner-Ellis theorem

In this section we discuss a different approach at small-time moderate deviations. While yielding only first order results, its conditions are usually easy to check for models with explicit characteristic function. Assumptions 2.1 and 2.2 are not in force here.

Recall that, in the classical setting of sequences of i.i.d. random variables, a moderate deviation analogue of Cramér's theorem can be deduced by applying the Gärtner-Ellis theorem to an appropriately rescaled sequence (see [9], Section 3.7). The MD short time behavior of diffusions can be subjected to a similar analysis. Consider the log-price  $X_t = \log S_t$  with  $X_0 = 0$  and mgf (moment generating function)

$$M(p, t) := \mathbb{E}[e^{pX_t}]. \quad (6.1)$$

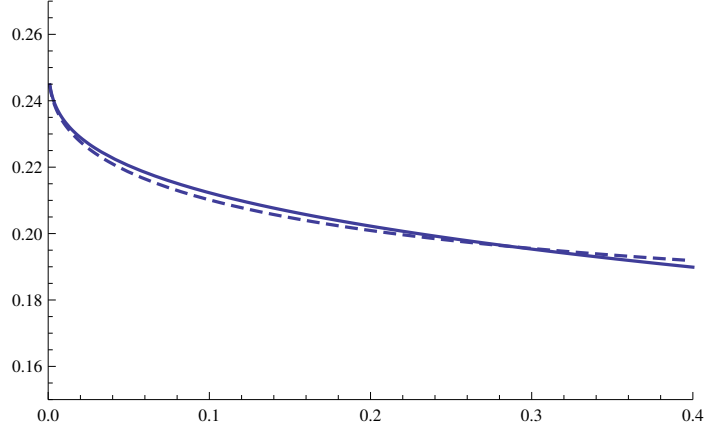


Figure 2: Illustration of our implied volatility expansion for the Heston model, with  $\ell \equiv \theta = 0.4$  and  $\beta = 0.3$ . Thus, log-strike equals  $k = 0.4t^{0.3}$ . The model parameters are  $\bar{v} = 0.0707$ ,  $\kappa = 0.6067$ ,  $\eta = 0.2928$ ,  $\rho = -0.7571$ ,  $v_0 = 0.0654$  (i.e.,  $\sigma_0 = 0.2557$ ), and  $S_0 = 1$ . The horizontal axis is time. The dashed curve is the exact MOTM implied volatility  $\sigma_{\text{imp}}(k_t, t)$ . The solid curve is the approximation  $\sigma_0 + \frac{\eta\rho}{4\sigma_0}k_t$  on the right hand side of (5.12).

**Assumption 6.1** For all  $\beta \in (0, \frac{1}{2})$ , the rescaled mgf satisfies

$$\lim_{t \downarrow 0} t^{1-2\beta} \log M(t^{\beta-1}p, t) = \frac{1}{2}\sigma_0^2 p^2, \quad p \in \mathbb{R}. \quad (6.2)$$

We expect that this assumption holds for diffusion models in considerable generality. It is easy to check that (6.2) holds for the Heston model, either by its explicit characteristic function, or, more elegantly, from the associated Riccati equations; see the forthcoming PhD thesis of A. Pinter for details. Thus, the results of the present section provide an alternative proof of the first order MOTM behavior (5.8) of Heston call prices.

Heuristically, Assumption 6.1 can be derived from the density asymptotics in Assumption 2.1, which in turn hold in quite general diffusion settings [10, 11]:

$$\begin{aligned} M(t^{\beta-1}p, t) &= \int e^{t^{\beta-1}px} q(x, t) dx \\ &\approx \int \exp\left(t^{\beta-1}px - \frac{\Lambda(x)}{t}\right) dx \end{aligned} \quad (6.3)$$

$$\approx \int \exp\left(t^{\beta-1}px - \frac{\Lambda''(0)x^2}{2t}\right) dx \quad (6.4)$$

$$= \exp\left(t^{\beta-1}px - \frac{x^2}{2\sigma_0^2 t} \Big|_{x=\sigma_0^2 p t^\beta} (1 + o(1))\right) \quad (6.5)$$

$$= \exp\left(\frac{1}{2}\sigma_0^2 p^2 t^{2\beta-1} (1 + o(1))\right), \quad t \downarrow 0. \quad (6.6)$$

In (6.3), we ignored that the density expansion (2.2) might not be valid globally in space; this might be made rigorous by estimating  $q(x, t)$  by a Freidlin-Wentzell LD argument for  $x$  sufficiently large. As for (6.4), we can expect concentration near  $x \approx 0$ , since  $\Lambda(x)$  increases with  $|x|$ . Finally, (6.5), and thus (6.6), follows from a (rigorous) application of the Laplace method. If (6.6) is correct, then (6.2) clearly follows.

The critical moment of  $S_t$  is defined by

$$p_+(t) := \sup\{p \geq 0 : M(p, t) < \infty\}.$$

It is obvious that

$$\lim_{t \downarrow 0} t^{1-\beta} p_+(t) = \infty \quad (6.7)$$

is necessary for (6.2), i.e.,  $p_+(t)$  must grow faster than  $t^{\beta-1}$  as  $t \downarrow 0$ . In the Heston model, e.g., the critical moment is of order  $p_+(t) \approx 1/t \gg t^{\beta-1}$  for small  $t$ , as follows from inverting (6.2) in [28]. On the other hand, we do not expect our results to be of much use in the presence of jumps. Indeed, suppose that (6.1) is the mgf of an exponential Lévy model. Then  $p_+(t) \equiv p_+$  does not depend on  $t$ , and is finite for most models used in practice. Therefore, (6.7) cannot hold, and so Assumption 6.1 is not satisfied. The Merton jump diffusion model is one of the few Lévy models of interest that have  $p_+ = \infty$ , but it is easy to check that it does not satisfy (6.2), either.

After this discussion of Assumption 6.1, we now give an asymptotic estimate for the distribution function of  $X_t$  (put differently, MOTM *digital call* prices) in Theorem 6.2. Then we translate this result to MOTM *call* prices in Theorem 6.3. If desired, higher order terms in (6.2) will give refined asymptotics in Theorem 6.2, using Gulisashvili and Teichmann's recent refinement of the Gärtner-Ellis theorem [25]. It might not be trivial to translate the resulting expansions into call price asymptotics, though. For other asymptotic results on option prices using the Gärtner-Ellis theorem, see, e.g., [14, 15].

**Theorem 6.2** *Under Assumption 6.1 (and without any further assumptions on our model), for  $k_t = \theta t^\beta$  with  $\beta \in (0, \frac{1}{2})$  and  $\theta > 0$ , we have a first order MD estimate for the cdf of  $X_t$ :*

$$\mathbb{P}[X_t \geq k_t] = \exp\left(-\frac{1}{2\sigma_0^2} \frac{k_t^2}{t} (1 + o(1))\right), \quad t \downarrow 0. \quad (6.8)$$

**Proof.** Define

$$Z_t := t^{-\beta} X_t, \quad \text{with mgf} \quad M_Z(s, t) = \mathbb{E}[e^{sZ_t}],$$

and

$$a_t := t^{1-2\beta} = o(1), \quad t \downarrow 0.$$

Then (6.2) is equivalent to

$$\Gamma_Z(p) := \lim_{t \downarrow 0} a_t \log M_Z(p/a_t, t) = \frac{1}{2} \sigma_0^2 p^2, \quad p \in \mathbb{R}.$$

Since  $\Gamma_Z$  is finite on  $\mathbb{R}$ , the Gärtner-Ellis theorem (Theorem 2.3.6 in [9]) implies that  $(Z_t)_{t \geq 0}$  satisfies an LDP (large deviation principle) as  $t \downarrow 0$ , with rate  $a_t$  and good rate function  $\Lambda_Z$ , the Legendre transform of  $\Gamma_Z$ . Trivially,  $\Lambda_Z$  is quadratic, too:

$$\begin{aligned} \Lambda_Z(x) &= \sup_{p \in \mathbb{R}} (px - \Gamma_Z(p)) \\ &= \sup_{p \in \mathbb{R}} (px - \frac{1}{2} \sigma_0^2 p^2) = \frac{x^2}{2\sigma_0^2}, \quad x \in \mathbb{R}. \end{aligned}$$

Now fix  $\theta > 0$ . Applying the lower estimate of the LDP to  $(\theta, \infty)$  yields

$$\begin{aligned} \liminf_{t \downarrow 0} a_t \log \mathbb{P}[Z_t \geq \theta] &\geq \liminf_{t \downarrow 0} a_t \log \mathbb{P}[Z_t > \theta] \\ &\geq -\Lambda_Z(\theta) = -\frac{\theta^2}{2\sigma_0^2}, \end{aligned}$$



and applying the upper estimate to  $[\theta, \infty)$  yields

$$\limsup_{t \downarrow 0} a_t \log \mathbb{P}[Z_t \geq \theta] \leq -\frac{\theta^2}{2\sigma_0^2},$$

and so

$$\lim_{t \downarrow 0} a_t \log \mathbb{P}[Z_t \geq \theta] = -\frac{\theta^2}{2\sigma_0^2}.$$

This is the same as (6.8). ■

As in the LD/OTM regime, first order cdf asymptotics translate readily into call price asymptotics. The proof of the following result is similar to [36], p. 30f (concerning the LD regime) and [6], Theorem 1.5. In the MD/MOTM regime, one can replace the condition (1.19) of [6] by a mild condition on the moments of the model.

**Theorem 6.3** *Let  $S = e^X$  be a continuous positive martingale. Assume that, for all  $p \geq 1$ , its  $p$ -th moment explodes at a positive time (infinity included). By this we mean that there is a positive  $t^*(p)$  such that the mgf  $\mathbb{E}[\exp(pX_t)]$  is finite for all  $t \in [0, t^*(p)]$ . Let  $v_0 = \sigma_0^2 > 0$ . Then the following are equivalent:*

(i) *For  $k_t = \ell(t)t^\beta$ , with  $\beta \in (0, \frac{1}{2})$  and  $\ell > 0$  slowly varying at zero, it holds that*

$$\mathbb{P}[X_t \geq k_t] = \exp\left(-\frac{1}{2v_0} \frac{k_t^2}{t} (1 + o(1))\right), \quad t \downarrow 0.$$

(ii) *Under the assumptions of (i), we have*

$$c(k_t, t) = \exp\left(-\frac{1}{2v_0} \frac{k_t^2}{t} (1 + o(1))\right), \quad t \downarrow 0. \quad (6.9)$$

**Proof.** First assume (i). Let  $\varepsilon > 0$  and define  $\tilde{k}_t = (1 + \varepsilon)k_t$ . Then

$$\begin{aligned} c(k_t, t) &\geq \mathbb{E}[(e^{X_t} - e^{k_t})^+ \mathbf{1}_{\{X_t \geq \tilde{k}_t\}}] \\ &\geq (e^{\tilde{k}_t} - e^{k_t})^+ \mathbb{P}[X_t \geq \tilde{k}_t]. \end{aligned} \quad (6.10)$$

The first factor is

$$(e^{\tilde{k}_t} - e^{k_t})^+ = (\tilde{k}_t - k_t + O(k_t^2))^+ = \varepsilon k_t + O(k_t^2).$$

For the second factor in (6.10), we apply (i) with  $\tilde{k}_t$ :

$$\lim_{t \downarrow 0} \frac{t}{\tilde{k}_t^2} \log \mathbb{P}[X_t \geq \tilde{k}_t] = -\frac{1}{2v_0}.$$

Therefore,

$$\begin{aligned} \liminf_{t \downarrow 0} \frac{t}{k_t^2} \log c(k_t, t) &\geq \lim_{t \downarrow 0} \frac{t}{k_t^2} \left(-\frac{1}{2v_0} \frac{\tilde{k}_t^2}{t} (1 + o(1))\right) \\ &= -\frac{(1 + \varepsilon)^2}{2v_0}. \end{aligned}$$

Now let  $\varepsilon \downarrow 0$  to get the desired lower bound for  $c(k_t, t)$ .

As for the upper bound, we let  $p > 1$  and note that, by definition of  $p \mapsto t^*(p)$ , we have  $\mathbb{E}[S_t^{p+1}] < \infty$  for all  $t \in [0, t^*(p+1)]$ . Define  $\bar{S}_t = \sup_{0 \leq u \leq t} S_u$  for  $t \geq 0$ . By Doob's inequality (Theorem 3.8 in [27]), we have

$$\mathbb{P}[\bar{S}_{t^*(p+1)} \geq s] \leq \frac{\mathbb{E}[S_{t^*(p+1)}^{p+1}]}{s^{p+1}}, \quad s > 0.$$

Hence  $\bar{S}_{t^*(p+1)}$  has a finite  $p$ th moment:

$$\mathbb{E}[(\bar{S}_{t^*(p+1)})^p] = p \int_0^\infty s^{p-1} \mathbb{P}[\bar{S}_{t^*(p+1)} \geq s] ds < \infty.$$

By the dominated convergence theorem and the continuity of  $S$ , we thus conclude

$$\lim_{t \downarrow 0} \mathbb{E}[S_t^p] = S_0^p. \quad (6.11)$$

Now let  $1/p + 1/q = 1$  and apply Hölder's inequality:

$$\begin{aligned} c(k_t, t) &= \mathbb{E}[(e^{X_t} - e^{k_t})^+ \mathbf{1}_{\{X_t \geq k_t\}}] \\ &\leq \mathbb{E}[(e^{X_t} - e^{k_t})^+]^{1/p} \mathbb{P}[X_t \geq k_t]^{1/q} \\ &\leq \mathbb{E}[S_t^p]^{1/p} \mathbb{P}[X_t \geq k_t]^{1/q}. \end{aligned}$$

By (6.11) and (i), we obtain

$$\limsup_{t \downarrow 0} \frac{t}{k_t^2} \log c(k_t, t) \leq \frac{1}{q} \limsup_{t \downarrow 0} \frac{t}{k_t^2} \log \mathbb{P}[X_t \geq k_t] = -\frac{1}{2qv_0}.$$

Now let  $p \uparrow \infty$ , i.e.,  $q \downarrow 1$ . The same argument yields the lower bound of the implication (ii)  $\implies$  (i). The remaining upper bound of (ii)  $\implies$  (i) is shown very similarly to the lower bound of the implication (i)  $\implies$  (ii). ■

## References

- [1] A. BENTATA AND R. CONT, *Short-time asymptotics for marginal distributions of semimartingales*. Preprint, available at <http://arxiv.org/abs/1202.1302>, 2012.
- [2] H. BERESTYCKI, J. BUSCA, AND I. FLORENT, *Asymptotics and calibration of local volatility models*, Quant. Finance, 2 (2002), pp. 61–69. Special issue on volatility modelling.
- [3] H. BERESTYCKI, J. BUSCA, AND I. FLORENT, *Computing the implied volatility in stochastic volatility models*, Comm. Pure Appl. Math., 57 (2004), pp. 1352–1373.
- [4] N. H. BINGHAM, C. M. GOLDIE, AND J. L. TEUGELS, *Regular variation*, vol. 27 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1987.
- [5] G. BRUNICK AND S. SHREVE, *Mimicking an Itô process by a solution of a stochastic differential equation*, Ann. Appl. Probab., 23 (2013), pp. 1584–1628.
- [6] F. CARAVENNA AND J. CORBETTA, *General smile asymptotics with bounded maturity*. Preprint, available at <http://arxiv.org/abs/1411.1624>, 2014.

- [7] P. CARR AND L. WU, *What type of process underlies options? A simple robust test*, The Journal of Finance, 58 (2003), pp. 2581–2610.
- [8] S. DE MARCO, P. FRIZ, AND S. GERHOLD, *Rational shapes of local volatility*, Risk, 2 (2013), pp. 82–87.
- [9] A. DEMBO AND O. ZEITOUNI, *Large deviations techniques and applications*, vol. 38 of Stochastic Modelling and Applied Probability, Springer-Verlag, New York, second ed., 1998.
- [10] J. D. DEUSCHEL, P. K. FRIZ, A. JACQUIER, AND S. VIOLANTE, *Marginal density expansions for diffusions and stochastic volatility I: Theoretical foundations*, Comm. Pure Appl. Math., 67 (2014), pp. 40–82.
- [11] ———, *Marginal density expansions for diffusions and stochastic volatility II: applications*, Comm. Pure Appl. Math., 67 (2014), pp. 321–350.
- [12] V. DURRLEMAN, *From implied to spot volatilities*, PhD thesis, Princeton University, 2004.
- [13] ———, *From implied to spot volatilities*, Finance Stoch., 14 (2010), pp. 157–177.
- [14] M. FORDE AND A. JACQUIER, *Small-time asymptotics for implied volatility under the Heston model*, Int. J. Theor. Appl. Finance, 12 (2009), pp. 861–876.
- [15] ———, *Small-time asymptotics for an uncorrelated local-stochastic volatility model*, Appl. Math. Finance, 18 (2011), pp. 517–535.
- [16] M. FORDE, A. JACQUIER, AND R. LEE, *The small-time smile and term structure of implied volatility under the Heston model*, SIAM J. Financial Math., 3 (2012), pp. 690–708.
- [17] M. FORDE AND H. ZHANG, *Asymptotics for rough stochastic volatility models*. Preprint, 2016.
- [18] P. FRIZ AND S. GERHOLD, *Extrapolation analytics for Dupire’s local volatility*, in Large Deviations and Asymptotic Methods in Finance, vol. 110 of Springer Proc. Math. Stat., Springer, Cham, 2015, pp. 273–286.
- [19] P. K. FRIZ, S. GERHOLD, AND M. YOR, *How to make Dupire’s local volatility work with jumps*, Quant. Finance, 14 (2014), pp. 1327–1331.
- [20] K. GAO AND R. LEE, *Asymptotics of implied volatility to arbitrary order*, Finance Stoch., 18 (2014), pp. 349–392.
- [21] J. GATHERAL, *The Volatility Surface, A Practitioner’s Guide*, Wiley, 2006.
- [22] J. GATHERAL, E. P. HSU, P. LAURENCE, C. OUYANG, AND T.-H. WANG, *Asymptotics of implied volatility in local volatility models*, Math. Finance, 22 (2012), pp. 591–620.
- [23] H. GUENNOUN, A. JACQUIER, AND P. ROOME, *Asymptotic behaviour of the fractional Heston model*. Preprint, available at <http://arxiv.org/abs/1411.7653>, 2014.
- [24] A. GUILLIN, *Averaging principle of SDE with small diffusion: moderate deviations*, Ann. Probab., 31 (2003), pp. 413–443.
- [25] A. GULISASHVILI AND J. TEICHMANN, *The Gärtner-Ellis theorem, homogenization, and affine processes*, in Large Deviations and Asymptotic Methods in Finance, vol. 110 of Springer Proc. Math. Stat., Springer, Cham, 2015, pp. 287–320.

- [26] P. HENRY-LABORDÈRE, *Analysis, geometry, and modeling in finance*, Chapman & Hall/CRC Financial Mathematics Series, CRC Press, Boca Raton, FL, 2009.
- [27] I. KARATZAS AND S. E. SHREVE, *Brownian motion and stochastic calculus*, vol. 113 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1991.
- [28] M. KELLER-RESSEL, *Moment explosions and long-term behavior of affine stochastic volatility models*, Math. Finance, 21 (2011), pp. 73–98.
- [29] A. L. LEWIS, *Option valuation under stochastic volatility*, Finance Press, Newport Beach, CA, 2000.
- [30] A. MIJATOVIĆ AND P. TANKOV, *A new look at short-term implied volatility in asset price models with jumps*. Preprint, available at <http://arxiv.org/abs/1207.0843>, 2012.
- [31] J. MUHLE-KARBE AND M. NUTZ, *Small-time asymptotics of option prices and first absolute moments*, Journal of Applied Probability, 48 (2011), pp. 1003–1020.
- [32] F. W. J. OLVER, *Asymptotics and special functions*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974.
- [33] Y. OSAJIMA, *General asymptotics of Wiener functionals and application to implied volatilities*, in Large Deviations and Asymptotic Methods in Finance, vol. 110 of Springer Proc. Math. Stat., Springer, Cham, 2015, pp. 137–173.
- [34] S. PAGLIARANI AND A. PASCUCCI, *The parabolic Taylor formula of the implied volatility*. Preprint, available at <http://arxiv.org/abs/1510.06084>, 2016.
- [35] L. PAULOT, *Asymptotic implied volatility at the second order with application to the SABR model*, in Large Deviations and Asymptotic Methods in Finance, vol. 110 of Springer Proc. Math. Stat., Springer, Cham, 2015, pp. 37–69.
- [36] H. PHAM, *Large deviations in finance*. Lecture notes for the third SMAI European Summer School in Financial Mathematics, Paris, August 2010.
- [37] S. R. S. VARADHAN, *Diffusion processes in a small time interval*, Comm. Pure Appl. Math., 20 (1967), pp. 659–685.
- [38] K. YOSIDA, *On the fundamental solution of the parabolic equation in a Riemannian space*, Osaka Math. J., 5 (1953), pp. 65–74.